Three-dimensional light-field microendoscopy with a GRIN lens array

TARA M. URNER,[1,2] Andrew Inman,[1,2] Benjamin Lapid,[1,2] and Shu Jia[1,*]

[1]*Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, 30332, USA*
[2]*These authors contributed equally to this work*
[*]*shu.jia@gatech.edu*

Abstract: Optical endoscopy has emerged as an indispensable clinical tool for modern minimally invasive surgery. Most systems primarily capture a 2D projection of the 3D surgical field. Currently available 3D endoscopes can restore stereoscopic vision directly by projecting laterally shifted views of the operating field to each eye through 3D glasses. These tools provide surgeons with informative 3D visualizations, but they do not enable quantitative volumetric rendering of tissue. Therefore, advanced tools are desired to quantify tissue tomography for high precision microsurgery or medical robotics. Light-field imaging suggests itself as a promising solution to the challenge. The approach can capture both the spatial and angular information of optical signals, permitting the computational synthesis of the 3D volume of an object. In this work, we present GRIN lens array microendoscopy (GLAM), a single-shot, full-color, and quantitative 3D microendoscopy system. GLAM contains integrated fiber optics for illumination and a GRIN lens array to capture the reflected light field. The system exhibits a 3D resolution of ~100 μm over an imaging depth of ~22 mm and field of view up to 1 cm^2. GLAM maintains a small form factor consistent with the clinically desirable design, making the system readily translatable to a clinical prototype.

1. Introduction

Physicians are accustomed to endoscopes that provide two-dimensional images originating from three-dimensional structures. However, surgeons' access to topological information in the axial dimension of the surgical field facilitated by rigid 3D endoscopy has been reported to reduce operation times and errors, especially in surgeries involving significant in-depth extension or 3D tissue complexity [1–9]. Furthermore, 3D endoscopy has shown promise for improving the learning experience of medical trainees [10–12].

Current clinical technologies incorporate stereoscopic imaging principles, in which two apertures record the tissue landscape simultaneously. The surgeon relays this information as a 2D image on either a head-mounted display or a specialized monitor that can be converted to a 3D perception using polarized eyewear [1]. While demonstrating the power of restoring depth perception to minimally invasive surgery, stereoscopic approaches suffer ergonomic and analytical downsides. Practically, the eyewear can cause dizziness and headaches after long periods of use, with some surgeons reporting excessive strain with the stereoscopic systems compared to conventional 2D endoscopy, even though operation times can be reduced [3,13]. Additionally, stereoscopic vision lacks quantitative 3D recording and reconstruction for intraoperative decisions, subsequent diagnostics, or data storage. As a result, surgeries may require other imaging procedures such as micro-CT or MRI to quantify the 3D morphology of tissue [14,15]. Indeed, acquiring quantitative volumetric information during surgical procedures has significant implications for diagnostics, treatment, and integration with digital and robotic devices.

In contrast, to achieve 3D reconstruction without eyewear, computational approaches to stereoscopic endoscopy such as deformable shape-from-motion and shape-from-shading have been proposed to quantify the 3D surface [16]. However, these algorithms are highly sample-dependent and may suffer from reduced temporal resolution due to required probe translation. Other quantitative approaches to stereoscopic imaging using epipolar geometry and the pinhole camera model have been shown to attain quantitative results but remain limited by the nonuniform field of view due to the use of only two apertures [17,18].

Light-field imaging is an optical methodology that addresses the limitations of the two-aperture approach in stereoscopic imaging. Light-field imaging, often used in 3D microscopy, is characterized by sampling the 2D spatial and 2D angular components of the light field with a lens or camera array [19–21]. For example, this approach has been applied to endoscopy by Orth and colleagues, who demonstrated that 3D light-field imaging could be obtained with a flexible multi-mode fiber bundle [22]. In contrast, rigid light-field endoscopy and otoscopy systems have also been proposed, utilizing microlenses or tunable electro-optic lens arrays, both of which, however, lead to a significant reduction in lateral resolution [23–28]. Therefore, there remains a demand for lens-based 3D endoscopy techniques that maintain high spatial resolution while providing quantitative depth information.

In this work, we demonstrate fast, 3D, multi-color microendoscopic imaging achieved by using a hexagonal gradient index (GRIN) lens array. This GRIN lens array microendoscopy system, or GLAM, provides a quantitative 3D imaging methodology with a high lateral and axial resolution. With the capability to detect the depth of features with sub-millimeter accuracy, GLAM is designed to meet many of the functional and physical requirements of a practical endoscope, including a small-diameter stainless steel shaft, built-in illumination, and multi-color imaging. With this combination of optical functionality and realistic endoscope design, GLAM demonstrates that quantitative 3D light-field imaging using GRIN lenses can be practically applied to rigid endoscopy, paving the way for future development using this optical paradigm. We expect GLAM to provide a necessary prototype to increase operative safety and efficiency with further implications on improving instrument control during robotic surgery.

2. Methods

2.1. Device instrumentation

In this study, we used a hexagonal GRIN lens array (GLA) to harness the benefits of light-field over stereoscopic imaging while mitigating the imaging quality trade-offs induced with conventional light-field methods (Fig. 1). Due to their optical properties, GRIN lenses can maintain a high numerical aperture (NA) within a physically confined space, allowing for dense 2D spatial sampling [29]. In this work, we incorporated our previous optical proof of concept of GRIN lens array imaging [30] into a handheld, compact, full-color realization with integrated illumination. Using 1-mm diameter GRIN lenses, we acquired 2D angular information at a pitch of 1.4 mm, keeping the entire probe diameter under 5 mm, consistent with 3D endoscopes currently used in the clinic [6] (Fig. 1(a)). In this sense, one 2D camera frame, consisting of seven angular elemental samples, one on-axis and six off-axis, permits a dense sampling of the light field to produce a quantitative 3D reconstruction. The GLAM prototype also contains six integrated optical fibers between the GRIN lenses for onboard illumination (Fig. 1(a),(b)). Additionally, we integrated full-color acquisition, gaining a more natural vision for better practicality (Fig. 1(a),(b)).

The detailed design and cross-sectional views can be found in Appendix 1 and Fig. 8. In brief, GLAM is constructed from a combination of off-the-shelf components and custom 3D printed parts. All custom parts were designed in SOLIDWORKS and printed with a resin 3D printer (Form 2, FLGPBK023D, Formlabs). The GLAM imaging probe is composed of an inner 3D printed core that houses seven 1-mm diameter 0.5NA GRIN lens assemblies (GRINTECH, GT-ERLS-100-005-150) arranged with a pitch of 1.4 mm hexagonally within individual tubes.
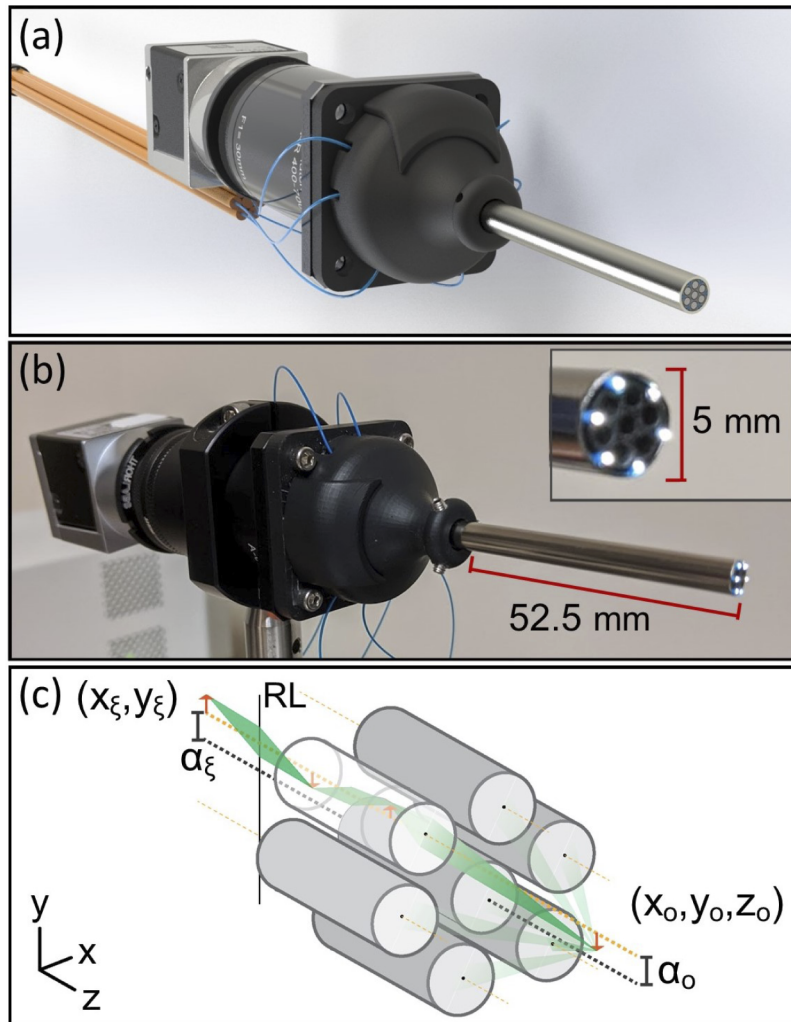
Fig. 1. GRIN lens array microendoscopy (GLAM) design and image formation. (a) CAD model of the endoscope assembly. (b) Assembled endoscope system with overlaid dimensions of the insertion tube. Inset shows a close-up of the 3D printed core of the insertion tube that holds the seven GRIN lenses and six illumination fibers. (c) Schematic diagram of light propagation through the system. RL, doublet relay lenses.

The GRIN lens assembly consists of a plane surface GRIN lens objective with a working distance of 5mm fused to a 15mm rod lens. Illumination is provided by an LED source (Thorlabs, MWWHF2), and the outside of the core has six grooves for threading optical fibers (Thorlabs, BF72HS01) between the inner core and outer sheath. The outer sheath is a 304 Stainless Steel Tube (McMaster-Carr, 8987K7) with an outer diameter of 0.203" (5.15 mm) and a wall thickness of 0.01" (0.25 mm). The probe fits into a custom holder with a side access door to align the lenses (Fig. 1(a),(b) and 8). The fibers were threaded through isolated channels in the holder, beyond which the coated GRIN lenses were extended to relay the image without significant interference from the light leakage. The image is relayed by an achromatic lens pair (Thorlabs, mAP103040-A) to a color camera (Basler Ace 304 acA1920-25uc).

2.2. *Algorithmic framework*

We utilize a ray-optics reconstruction without any computationally costly deconvolution to reduce the computational burden for eventual clinical application [31]. This approach enables full-color quantitative reconstruction of spatially dense samples over centimeters of depth while reducing reconstruction time by more than an order of magnitude over the previous wave-optics model [30]. The current reconstruction time at a speed of ~0.9 seconds per millimeter can be further increased when integrated with the GPU architecture. The reconstruction procedure is based on the optical parametrization outlined in Fig. 1(c). Each lens samples a differently angled cone of light rays reflected from each point on the object. The object space coordinates are denoted as (x_o, y_o, z_o), and the image space coordinates are denoted as (x_ξ, y_ξ). GRIN lenses are spaced with a pitch of α_o in object space and correspondingly, α_ξ in image space. The ray optics formulation of the light-field reconstruction can be derived as a shearing process in 3D space [27,32]. For GLAM, the mapping of the object space to the image space over an axial depth z is given by the depth-dependent magnification function $M(z)$. By pre-calibrating the magnification of the system, each layer of the 3D reconstruction can be encoded with a quantitative depth. A detailed discussion of the point-spread function (PSF) calibration can be found in Appendix 2. The latest version of the software is available in Ref. [31].

3. Results

3.1. *Chromatic characterization*

GLAM is calibrated through the acquisition of the PSF of the system. A pinhole is aligned along the optical axis of the center GRIN lens and translated axially. Figure 2(a) shows the schematic
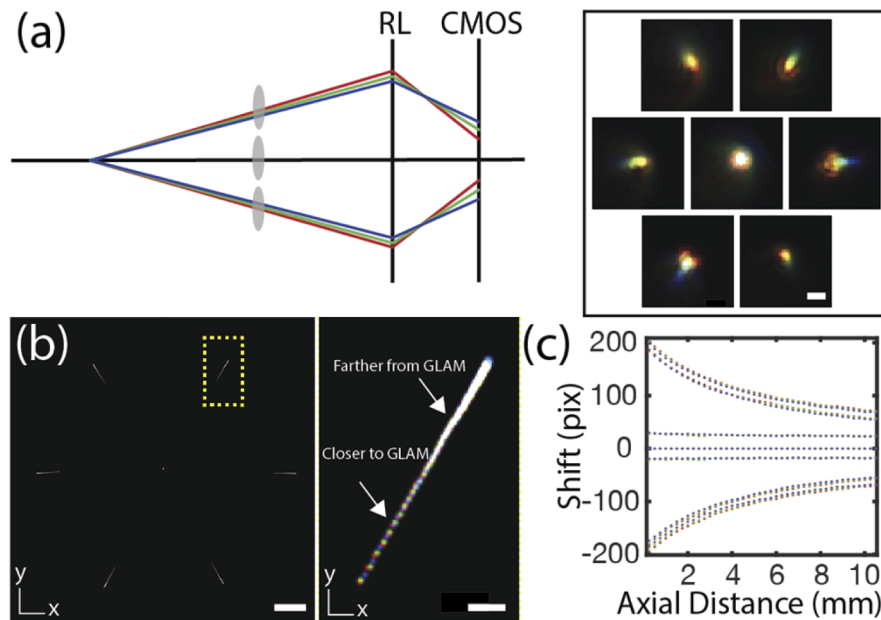

Fig. 2. Chromatic calibration of GLAM. (a) Schematic representation of axial chromatic aberration in the system. Inset shows laterally displaced RGB PSF positions in the off-axis elemental images. (b) Axial stack projection (step size = 50 μm) of the RGB PSF within an axial range of 22 mm. Inset shows the zoomed-in image of the yellow boxed region. The colors can be clearly resolved closer to GLAM, indicating an enhanced axial resolution. (c) Projection of the PSF shift from lens center for each elemental image along the *y-z* axis. A 0-10 mm axial displacement range is shown for better visualization of the point separation close to the endoscope. Scale bars: 20 μm (a), 500 μm (b), 50 μm (b, inset).
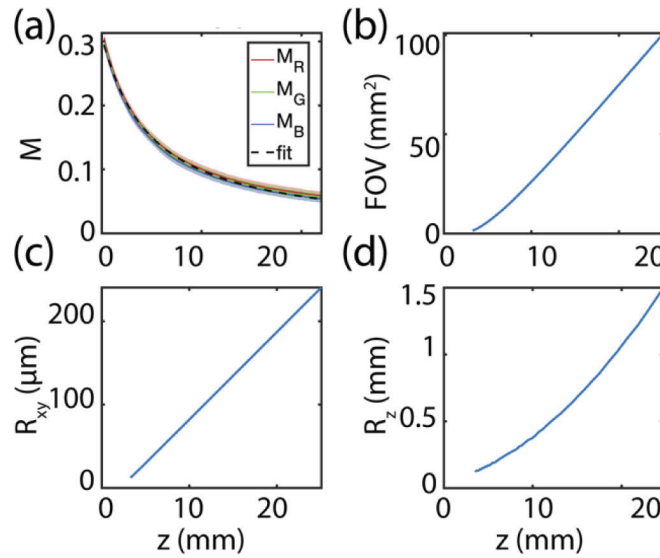
Fig. 3. System characterization using *M*(z). (a) Magnification of RGB as a function of the distance of the object from the system. Solid lines show the average magnification for each color, and the shaded areas represent the standard deviation over all the lenses. (b) FOV, defined as the overlapping region of all seven lenses in the reconstruction. (c) The Nyquist sampling resolution limit of the system. (d) Axial resolution given as the smallest resolvable axial shift over distance.

depiction of the RGB acquisition and one axial plane of the resulting RGB PSF. Chromatically dependent displacement can be observed in the six off-axis GRIN lenses, with longer wavelengths tending towards the center of the GLA and shorter wavelengths shifted outward from the center. As the off-axis lenses capture angular information of the pinhole on the central optical axis, this displacement is a direct measurement of the axial chromatic effect in the system, an aberration common to GRIN lenses [33]. As the pinhole is translated towards the system, the chromatic shifts change as a function of the magnification, as shown in Fig. 2(b), displaying an *x-y* projection of the entire 3D PSF over an axial range of 22 mm. It can be seen that the RGB pinhole images are more spread apart when the pinhole is close to the system and gradually come closer together as moved away. The corresponding pixel shifts are quantified as in Fig. 2(c), where the shift value is projected along the *y-z* axis. With the ray-optics model described in Appendix 2, we determine the axial offset between the RGB channel images to be ~36 μm, reasonably close to the nominal value of ~24 μm provided by the manufacturer for the GRIN-relay system alone.

3.2. Theoretical resolution and field of view

The RGB magnification curves were averaged prior to fitting the model to determine the average pinhole position, shown in Fig. 3(a). The magnification function of the system determines the pixel size of the image space, which sets the upper limit on the field of view (FOV) and sampling resolution (Fig. 3(b),(c)). The FOV is a function of axial distance and is defined here as the area where all seven lenses contribute intensity information within each reconstructed axial plane. Such an area exhibits the highest SNR for the final image, though regions outside this FOV can still contribute to the reconstruction [30]. Individual RGB lens *M*(z) data and experimental measurements of the FOV are shown in Appendix 3. With our model, the sampling resolution remains under 250 μm and the FOV can achieve 1 cm² at 20 mm away from the system (Fig. 3(b),(c)). The sampling resolution becomes worse due to the effects of diffraction and

aberrations present in the system. The axial resolution limit of GLAM can be conceptualized in opaque samples as the smallest axial shift that translates to a measurable lateral shift on the camera. Figure 3(d) shows the theoretical limit for axial translations or topological features that can be resolved with GLAM.

3.3. Characterization of lateral resolution

Experimentally, the axially-dependent magnification of the system within a range of 0-22 mm from the endoscope was determined using the PSF (Fig. 4(a)). The magnification can also be used to calculate the effective image pixel size at different depths, which were used to estimate the resolution of the system. To quantify the lateral resolution of GLAM, we mounted and imaged a USAF resolution target (R1DS1N, Thorlabs) on a motorized linear translation stage. These images of the target were taken using transmitted light. Figure 4(b) shows a fused image of the RGB intensity values recorded from 2.85 mm away, and the red line indicates element 5 of group 4 on the target as the finest resolvable element. The intensity graph along the red line is shown in Fig. 4(c), where the three bars of the element were identified using Gaussian fitting. The distance between peaks was used to estimate the resolution at that depth, and the data for fused RGB images at all measured depths are shown in Fig. 4(d). The lateral resolution
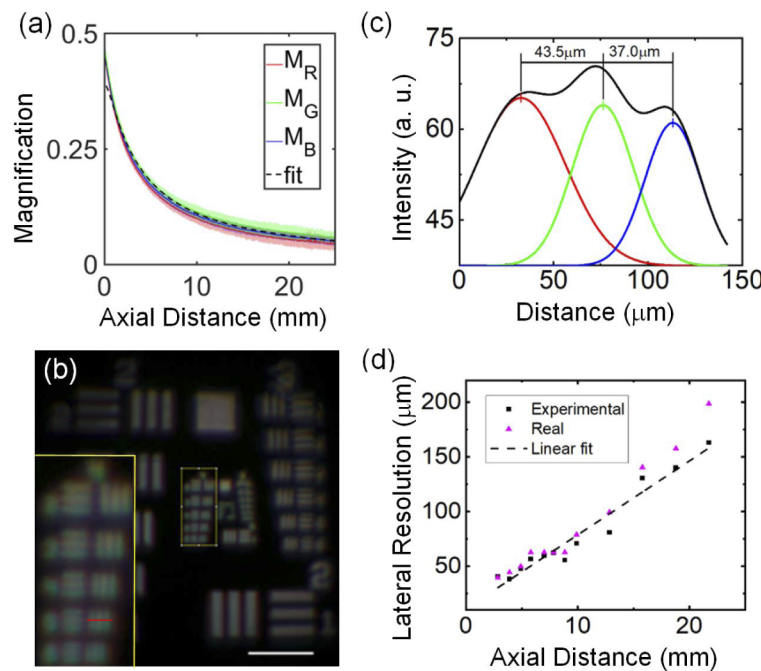

Fig. 4. Experimental system characterization. (a) Magnification of RGB as a function of distance from the system. Solid lines show the average magnification for each color, and the shaded areas represent the standard deviation over all the lenses. (b) Reconstructed RGB images of the USAF target taken at 2.85 mm away from GLAM. Inset shows the zoomed-in image of Group 4 (physical size: $535\,\mu m \times 1173\,\mu m$). (c) The black curve shows the overall intensity profile along the finest resolvable elements indicated by the red line in (b). The red, green, and blue curves represent the Gaussian-fitting of the three corresponding bars. (d) Measurements of the experimental lateral resolution obtained for RGB data as a function of distance from GLAM. The magenta points represent the real distance between the bars as a function of distance from GLAM, and the dashed line represents the line of best fit for the data. Scale bar: 1 mm.

exhibits a linear trend, extending from 38 µm to 162 µm across a depth range of 1.85-21.75 mm away from the endoscope, consistent with the predicted results in Fig. 3(c). Due to chromatic
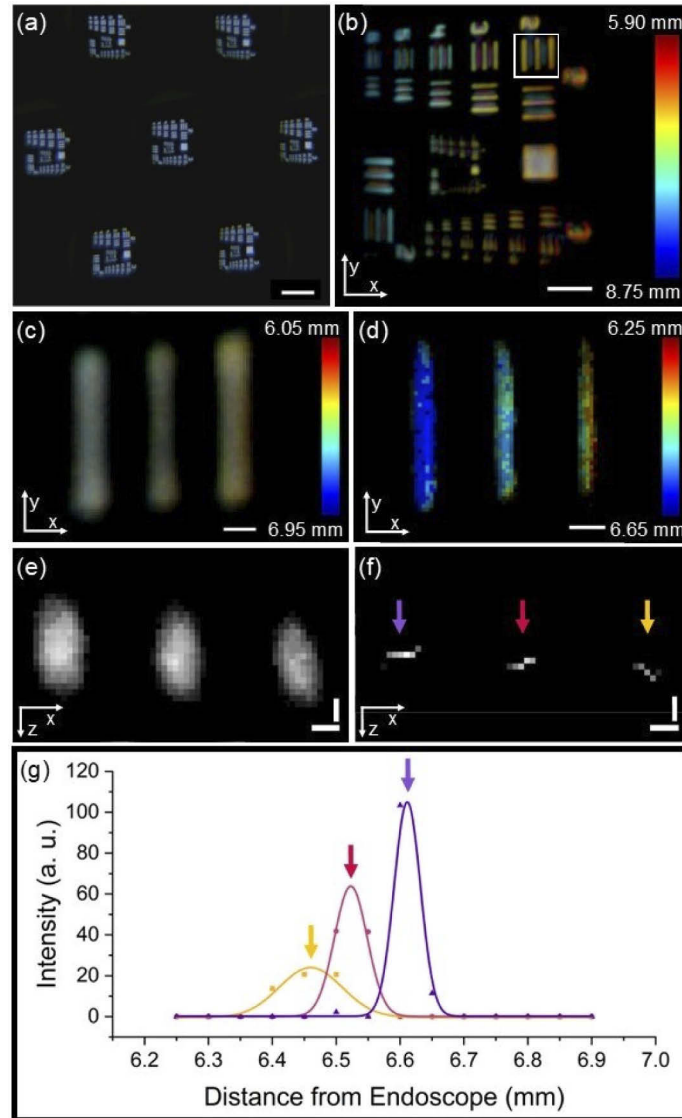

Fig. 5. Axial resolution measurement. (a) Raw, full-color endoscope image of a USAF resolution target angled at 20°. (b-d) Full-color reconstructions with true color overlaid with a color gradient to show depth. (b) Full-color reconstruction of the angled USAF target. (c) Full-color reconstruction of the (2, 2) bars, white boxed in (b). (d) Reconstruction of the (2, 2) set of bars after isolating the weighted maximum intensity, exhibiting the clear depth gradient. (e) Projected top view of the reconstruction in (c). (f) Projected top view of the reconstruction in (d). (g) Intensity profiles of the projected bars in (f). Gaussian fitting gives centers of the intensity profile of each bar at 6.460 mm, 6.523 mm, and 6.611 mm from the tip of the endoscope, showing a resolved 88-µm distance between the bars with centers at 6.523 mm and 6.611 mm. Scale bars: 500 µm (a), 750 µm (b), 100 µm (c, d), 50 µm (e, f, vertical), 200 µm (e, f, horizontal).

aberration, the actual distances between element bars of the USAF target group are slightly larger than the experimental values. The detailed data for characterizing the lateral resolution for each color can be found in Appendix 4 and Fig. 14.

3.4. Characterization of axial resolution

The axial resolution of the system was measured using the same USAF resolution target and transmitted light. The USAF target was mounted on a rotating stage, allowing imaging of the target at a range of angles. Angling the target introduced a variable deviation in the axial position of the bars on the target that was used to determine the smallest axial distance that the system could resolve. The target was imaged at 0°, 10°, 20°, 30°, 40°, and 45° angles. The distance between the target and the endoscope was such that the middle bar of the (2,2) group on the USAF target was 6.5 mm from the endoscope when the target was angled at 20° (Fig. 5). At each target angle, as shown in Fig. 5(a), the raw image was used to generate the full-color reconstructions shown in Fig. 5(b), which were color-coded according to the depth due to the angling of the target. The reconstructions were then processed by isolating the weighted maximum intensity of each pixel throughout the reconstruction stack to remove out-of-focus information in the z-axis (Fig. 5(c)-(f)). The intensity profiles of the bars in the z-axis of the top-view projection (x-z projection) were each fit to a Gaussian distribution (Fig. 5(g)). We measured that the system is able to resolve the axial distance between the two Gaussian curves separated by 88 µm, consistent with the theoretical distance of 76 µm for the adjacent bars on the USAF target angled at 20°.

3.5. Phantom curvature estimation

Next, to further assess the depth estimation ability of the GLAM system, we imaged a phantom target resembling red and blue blood vessels that were wrapped around a half-inch diameter tube. Figure 6(a) depicts the raw image of the target through the center GRIN lens of the endoscope. A color reconstruction was made from the data. The reconstruction was then converted to the grayscale intensity, and the intensities were inverted to aid in processing. For each pixel, a plane of focus was determined using the intensity and a weighted maximum approach. The planes
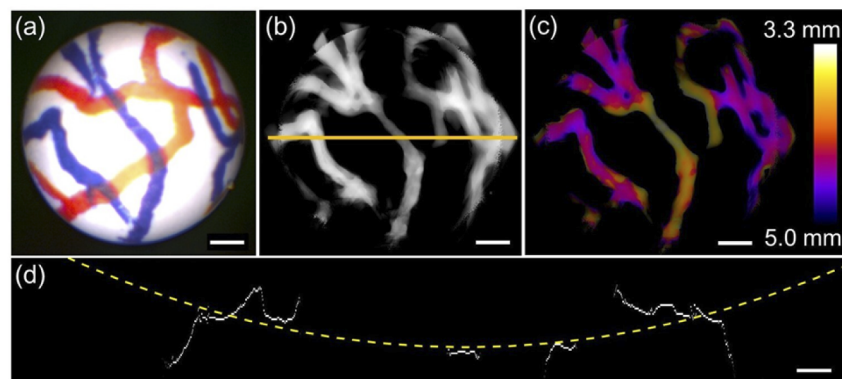

Fig. 6. Imaging phantom curvature. (a) Raw image from center endoscope lens of phantom blood vessels wrapped around a 0.5-inch diameter cylinder. (b) Full-color reconstruction of the phantom target, shown in an inverted grayscale image. The weighted maximum of each pixel in the reconstruction stack was extracted, and the resulting stack slices were projected into a single plane. (c) Depth-coded reconstruction in (b) with distance from the endoscope shown by a color gradient. (d) Projected top view of the reconstruction stack along the yellow line in (b). The profile was fitted with a dashed circle with a diameter of 0.546 inches. Scale bars: 1 mm (a-c). 500 µm (d).

of focus ranged between 3.04 mm and 5.99 mm away from the endoscope. Figure 6(b) shows a projection of all stack slices into a single plane. Figure 6(c) displays the reconstruction with a color gradient to represent the depth of the plane of focus for each pixel, consistent with the expected depth estimation where the center of the phantom target is closer to the endoscope compared to the outer edges of the target. Figure 6(d) shows a top-down view projection of the reconstruction in Fig. 6(b) along the yellow line. The sections of pixels that had a more horizontal trend were used to fit a circle to estimate the surface curvature of the target. The curvature rendered a 0.546-inch diameter, consistent with the actual 0.5-inch diameter of the cylinder.
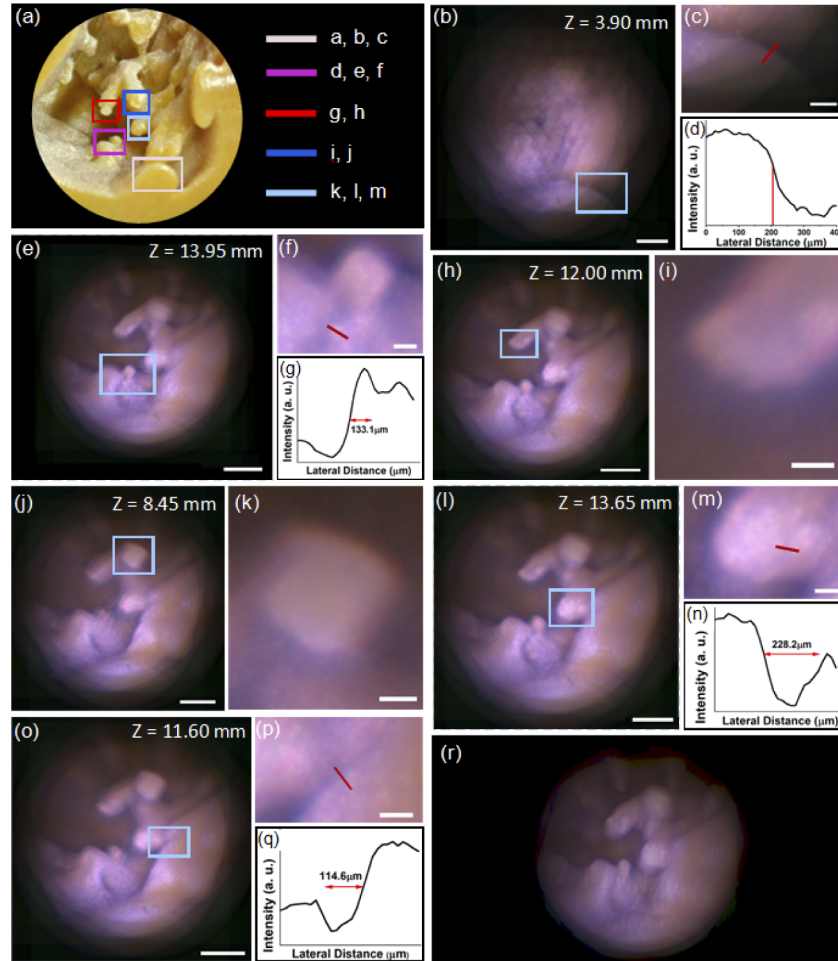

Fig. 7. Imaging phantom heart model. (a) 2D Image of the model with the features used in this figure marked. (b, e, h, k, l, o) Full field-of-view reconstruction slices at various depths $z = 3.90, 13.95, 12.00, 8.45, 13.65, 11.60$ mm from the tip of the endoscope, respectively. (c, f, i, k, m, p) Corresponding zoomed-in images of the boxed regions of (b, e, h, j, l, o), respectively. (d, g, n, q) Intensity profiles plotted along the red bars in (b, e, l, o), respectively. The red line in (d) indicates the region of the steepest intensity profile. (r) Volumetric rendering of the model. Scale bars: 1.5 mm (b), 4.0 mm (e), 3.5 mm (h), 2.5 mm (j), 3.0 mm (l, o), 0.5 mm (c, f, i, k, m, p).

3.6. Quantitative 3D reconstruction of a phantom heart

Lastly, we validated the performance of the GLAM system for phantom organs that contain fine features at various axial positions. In particular, the anatomical structures within a 3D printed heart model were imaged (Fig. 7). With the light-field acquisition, we reconstructed the volume of the model and generated the synthesized stack slices (e.g., Fig. 7(b),(c) at $z = 3.90$ mm from the tip of the endoscope). Here, we finely determined the depth of the features by locating the reconstruction stack slice with the steepest slope in the intensity plot (e.g., Fig. 7(d)), and a more detailed explanation for this process can be found in Appendix 5. For example, we selected several reconstructed features at the depths $z = 13.95$ mm (Fig. 7(e) and 7(f)), 12.00 mm (Fig. 7(h) and 7(i)), and 8.45 mm (Fig. 7(j) and 7(k)). In comparison, the digital caliper measurement of these features resulted in corresponding physical distances at 13.92 mm, 11.72 mm, and 8.26 mm, indicating good agreement of the quantitative depth rendering using GLAM. Caliper measurements were made relative to the flat face shown in Fig. 7(b),(c) and boxed in purple in Fig. 7(a). Furthermore, the structures of the heart model with feature sizes of 100-200 µm can be well resolved (Fig. 7(g), 7(n), and 7(q)), consistent with the prior calibrated lateral resolution of ~100 µm at these corresponding depths (Figs. 3–4). In Fig. 7(r). a volumetric view of the sample generated with a maximum brightness 3D projection algorithm in Fiji [34].

4. Discussion and conclusions

In summary, we demonstrate GLAM, a compact, single-shot, full-color, and quantitative 3D microendoscopy system. By subsampling the angular component of the light field, we gain access to the axial dimension, achieving a 3D resolution of ~100 µm over an imaging depth of ~22 mm and field of view up to ~1 cm^2. The system incorporates a GRIN lens array instead of the two-lens stereoscopic scheme, promising an alternative paradigm for clinical applications requiring high-resolution, quantitative volumetric measurements. GLAM exhibits a small form factor, a prototype readily translatable to further preclinical and clinical testing.

Specifically, in our approach, we utilize the system PSF for pre-calibration of the 3D reconstruction algorithm. This has the advantages of 1) considering misalignments or other experimental anomalies and 2) making the quantitative 3D reconstruction sample-independent. One current limitation of the PSF calibration approach is that it assumes a nominal lens separation of 1.4 mm. This may be slightly different between each lens due to inhomogeneities in the 3D printed core. Future versions of the analysis software could address this by calibrating the lens pitch directly to the GLAM system. Also, this may improve the estimation of the PSF offset and GRIN-to-relay spacing. For the reconstruction, speeds of ~0.9 seconds per millimeter have been obtained over multiple millimeters of depths without any further optimization of the algorithm or processing hardware. Through use of a GPU and optimization of the reconstruction algorithm to fully utilize this hardware, the reconstruction algorithm should be able to obtain video-rate real-time 3D imaging and visualization.

The use of the GLA enables quantitative depth estimation and allows for simple chromatic calibration for accurate RGB depth encoding. The pinhole image offsets in the off-axis elemental images provide a quick readout for axial chromatic aberrations in the GLA, which we have incorporated into the system magnification function $M(z)$. Traditionally, this information would be obtained through a more complicated imaging protocol involving scanning optics and a fluorescent sample. In contrast, this calibration method offers a fast readout of axial aberrations suitable for incorporating into the downstream analysis.

The 3D reconstruction obtained with the GLAM system demonstrates robust axial sectioning capability and shows recovered depth information about opaque, reflective samples on the microscale. Notably, the GLA assembly used in this system can be used as a blueprint that is readily reconfigurable and scalable by altering the pitch and focal distance of the system to match desired applications. Additionally, GRIN-to-relay spacing or additional relay lenses can tune the

magnification function as needed. Increasing the pitch will improve axial resolution with the trade-off of a smaller field of view and a more prominent form factor, while decreasing pitch has the opposite effect. An exemplary effect of adjusting GRIN-to-relay spacing can be seen in Fig. 10 in Appendix 2.

The use of integrated illumination is another characteristic that makes the endoscope system viable for practical use. Around the GRIN lens array are six optical fibers, which provide uniform illumination to the area in front of the endoscope. A computer can continuously control the intensities of the illumination to provide the appropriate amount of lighting for data sampling.

The properties, including the high 3D resolution, full-color acquisition, and computational simplicity, promise GLAM for future advancement and realization for surgical procedures. Furthermore, the system offers the potential to integrate quantitative 3D imaging with other devices such as surgical robotics to conduct more accurate automated 3D navigation within tight spaces in the body. Such a strategy for microimaging in three dimensions could be extended beyond the medical realm for general engineering and manufacturing purposes.

Appendices

Appendix 1: Additional system details

The simulated magnification shown in Fig. 3(a) was calculated through application of the thin lens approximation, $\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i}$, where d_o is the distance from the lens to the object, f is the focal length of the lens, and d_i is the distance from the back of the lens to the image. The focal length of the GRIN lenses was calculated as $f_g = \frac{1}{n_o g \sin(gl)}$, where n_0 is the center index of the GRIN lens, g is the gradient constant, and l is the length of the lens. These equations were applied along with the magnification equation, $M = \frac{d_i}{d_o}$, where d_i is the distance between the lens and the image and d_o is the distance between the lens and the object, to generate the magnification function for the system, $M(z)$, as shown below [see Eqs. (1)–(7)]. This magnification equation was then used to simulate the lateral and axial resolution of the system, as described in Appendix 3.

$$di_0 = \left(\frac{1}{f_g} - \frac{1}{z}\right)^{-1} \tag{1}$$

$$do_1 = d_{gtr} - di_0 \tag{2}$$

$$di_1 = \left(\frac{1}{f_{relay}} - \frac{1}{do_1}\right)^{-1} \tag{3}$$

$$do_2 = d_r - di_1 \tag{4}$$

$$di_2 = \left(\frac{1}{f_{relay}} - \frac{1}{do_2}\right)^{-1} \tag{5}$$

$$M_{System} = M_{GRIN} * M_{Relay\ 1} * M_{Relay\ 2} \tag{6}$$

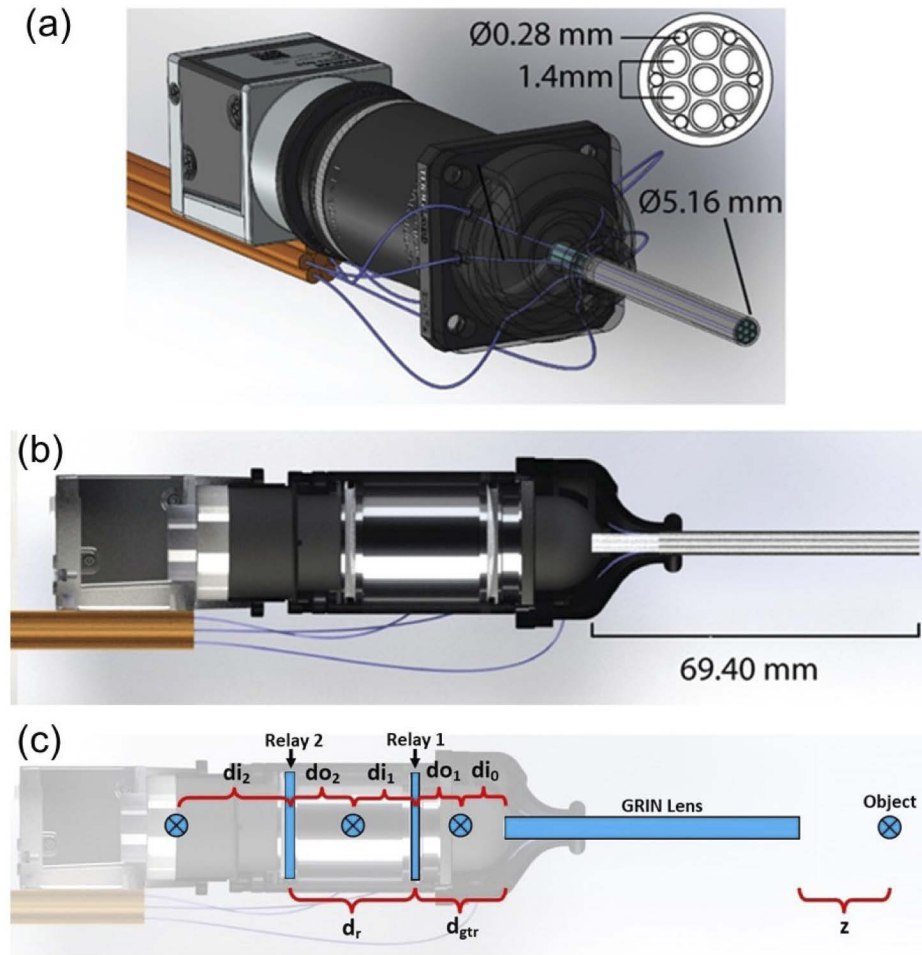$$M(z) = \frac{di_2}{do_2} * \frac{di_1}{do_1} * \frac{di_0}{z} \tag{7}$$

Fig. 8. Transparent and cutaway renderings of GLAM. (a) A CAD model of the GLAM showing fiber implementation into the imaging probes. Each fiber is individually threaded into the imaging probe. (b) Cutaway view of the GLAM system showing the GRIN lens extension past the fiber entry point followed by two achromatic doublets and the CMOS RGB camera. (c) Cutaway view of the GLAM system with highlighted lenses and distances used for magnification simulation.

Appendix 2: System calibration

A light-field system maps axial shifts in the object space to lateral shifts in the image space – thus encoding the depth into 2D information. The reconstruction process is an inverse mapping, i.e., a back projection. As a pre-calibration step for reconstruction, the PSF of the system is acquired as depicted schematically in the box within Fig. 9(a). First, a pinhole is aligned on the optical axis of the center lens and placed as close as possible to the system – displaced by an arbitrary offset O. Light from the pinhole passes first through each of the seven lenses, six of them offset from the center lens by a pitch α_0, and then through the achromatic lens pair to form the final image on the camera. The distance from the relay lens to the camera is a set parameter in our system, while the distance from each GRIN rod lens to the relay lens pair is tunable during alignment, thus represented by the distance variable a. The final magnification of the pinhole image can be seen in this case as $M = \frac{h_\xi}{h_o} = \frac{(x_\xi - \alpha_\xi)}{\alpha_o}$. As the pinhole is translated in steps of $\Delta z_0 = 50$ µm,

there is a corresponding shift in image space given by Δx_ε. The shift Δx_ε is proportional to the magnification change of the system over Δz_o, $M(\Delta z_o) = \frac{h_\xi}{h_o} = \frac{(\Delta x_\varepsilon - \alpha_\xi)}{\alpha_o}$. Figure 9(b) shows a cumulative z projection of the entire PSF volume. Though the pinhole remains on the optical axis in the center lens, and thus h_o is not changing, we can see that x_ε has changed for every axial shift according to the depth-dependent magnification of the system $M(z)$. Using the information provided by the PSF as a guide, the reconstruction procedure shifts and averages the seven sub images at each axial step. At each step, the images are overlapped such that each off-axis sub
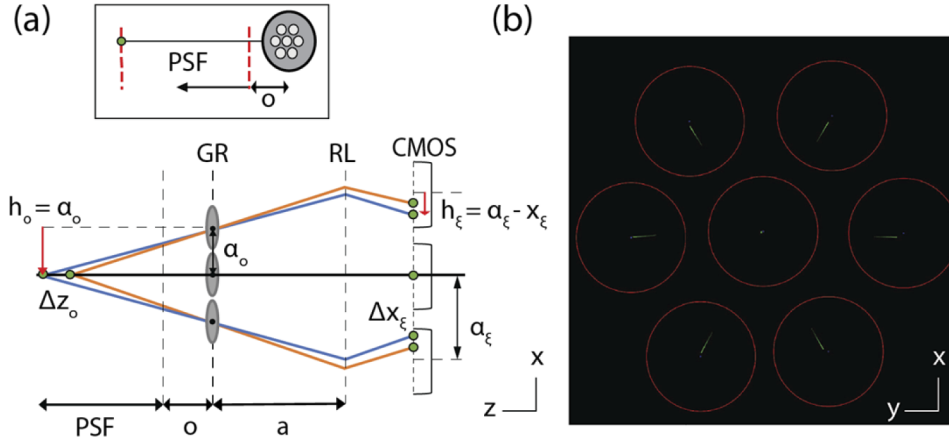

Fig. 9. Magnification model. (a) Two different positions of the pinhole are represented by the green dot. With each axial shift, Δz_0 is encoded in the PSF with a lateral shift Δx_ε. GR is the GRIN/relay system, RL is the achromatic doublet relay lens pair. The two variables O and a represent the offsets of the PSF region from the system and the GRIN-to-relay spacing, respectively. (b) An *x-y* projection of the PSF volume, showing the cumulative shifts in the pinhole image in the off-axis lenses, proportional to the magnification of the system.
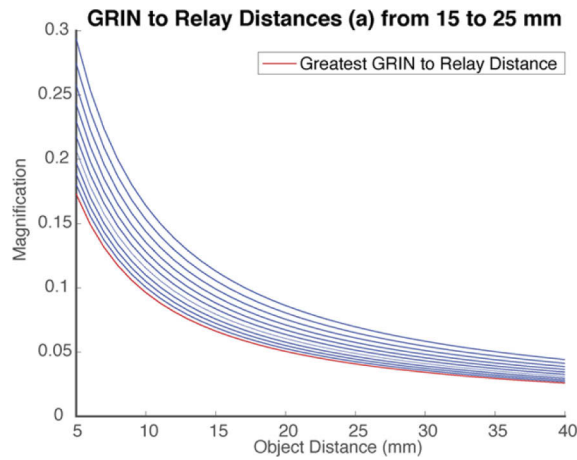

Fig. 10. Ray-optics simulations of $M(z)$. Simulations of changing GRIN-to-relay distance over a 40-mm axial distance. The farther the lenses are from each other, and the more gradual the magnification decrease will be. This corresponds to a larger FOV but lowers the axial and lateral resolution. A changing O parameter can be conceptualized as shifting this graph left or right for different offsets.

image of the pinhole would align with the center image, thus ensuring that each axial plane of the reconstruction is in focus.

To fully characterize the mapping from the object to image space, we further solved for the unknowns offset O and GRIN-to-relay spacing a. This is achieved by fitting the experimental $M(z)$ with a ray-optics model of magnification in our system. Changing the GRIN-to-relay spacing will alter the magnitude of $M(z)$, while different PSF offsets will left or right shift the portion of the curve captured by our PSF. Figure 10 shows example $M(z)$ functions with different GRIN to relay spacing. The experimental magnification data is fit to this model to simultaneously solve for both a and O.

Appendix 3: System alignment and characterization

Each of the seven GRIN lenses in GLAM is separately aligned until its image sharpness reaches a maximum. Characterization of the lens-by-lens alignment of the off-axis lenses can be seen in Fig. 11.
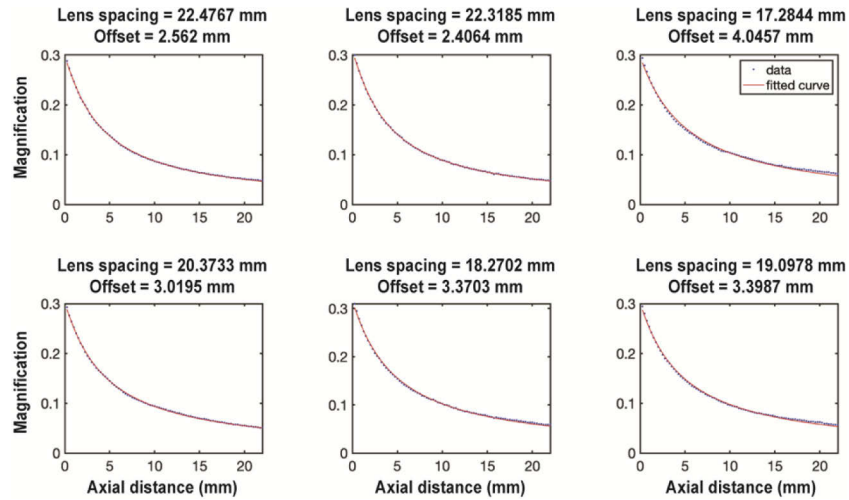

Fig. 11. Magnification function $M(z)$ per lens. Magnification fits for each of the seven lenses in a single channel. There are slight shifts in the lenses representing misalignments in the system. An extra iterative calibration step could be added to pre-calibration to minimize the difference in these fits before imaging.

The six $M(z)$ functions are averaged in each RGB channel, as shown in Fig. 12. The theoretical lateral and axial resolution limits shown in Fig. 3(c)-(d) are then derived from the average curve. The Nyquist sampling resolution limit for lateral resolution is given at each axial position by $R_{xy} = 2(M(z) \times S_{pixel})$, where S_{pixel} is the physical size of the camera pixel. The axial resolution of the system can be thought of as the smallest axial shift that will result in an observable lateral shift on the camera. This is calculated by finding the size of the axial shift around a given reference z position that results in a lateral shift greater than the lateral resolution limit on the camera $2(M(z \pm \Delta z) \times S_{pixel}) > R_{xy}$.

The effective field of view of GLAM will change over the imaging depth depending on the overlap in the viewing region of each lens as well as the changing effective pixel size Fig. 3(b). Figure 13 shows FOV images for three distances from the GLAM system. Closer to the system, there is less overlap from the off-axis lenses, reducing the FOV. Farther away from the system, the overlap and pixel size increase, resulting in an increased FOV.
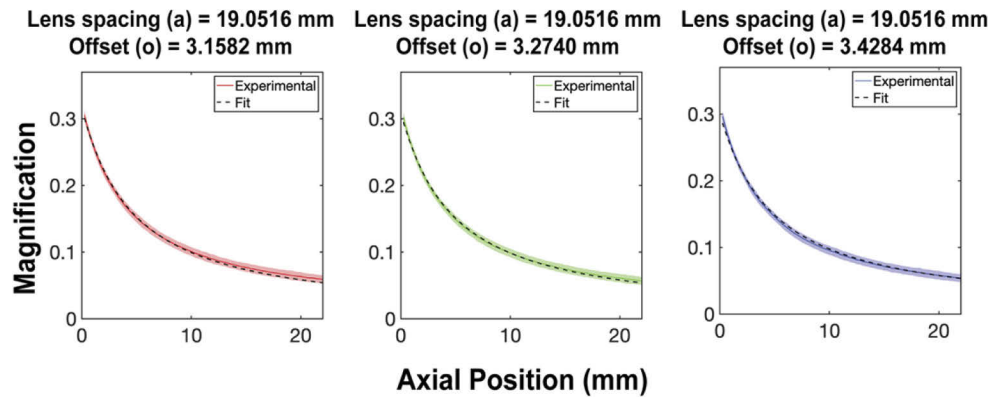
Fig. 12. RGB $M(z)$ with fixed lens spacing (a). As the RGB PSF is acquired for all the colors simultaneously, parameter (a) should not change between channels. The thick line shows the bounds of the individual lens fits, whose average is the solid line. The dotted line represents the fit to the theoretical model. When a is fixed, we can see the shifts in offset caused by axial chromatic aberrations in the system.
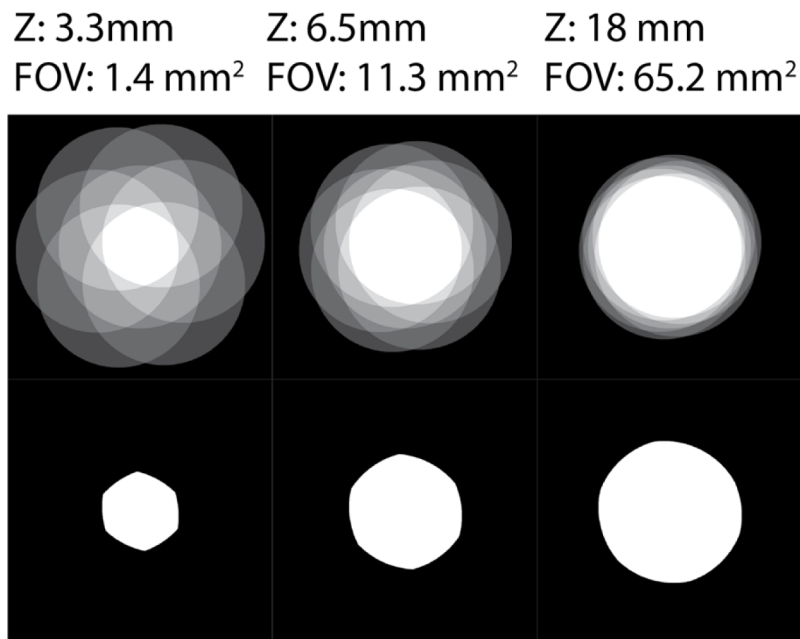

Fig. 13. FOV calibration. The field of view of GLAM was calibrated by reconstructing a binary mask of the GLA. The top row shows the FOV with all seven lenses, and the bottom region shows the area where their viewing regions overlap. The latter was considered the FOV, though other regions can contribute lower SNR information to the reconstruction.
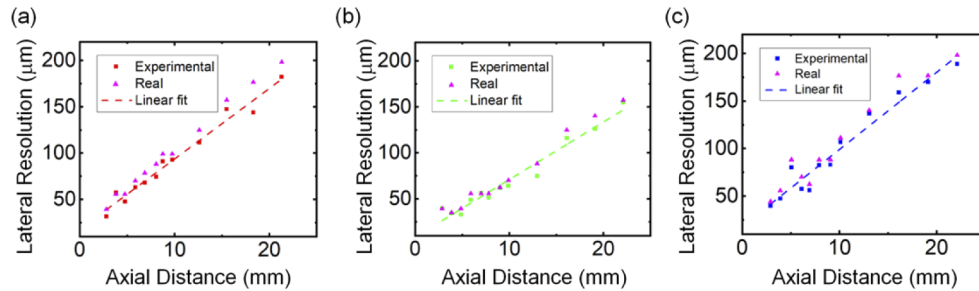
Appendix 4: RGB lateral resolution

Fig. 14. Measurements of the experimental lateral resolution obtained for red (a), green (b), and blue (c) intensity data as a function of distance from the system. The magenta points represent the real distance between the bars of the USAF target as a function of distance from the system, and the dashed lines represent the line of best fit for their respective data sets.

Appendix 5: Determination of axial focus plane

The axial distances of the heart model features shown in Fig. 7 were determined by first plotting the intensity of pixels along a line that crosses the edge of a feature. Figure 15 provides an example of this process for the feature shown in Fig. 7(j) and Fig. 7(k). A range of 15 stack slices was analyzed for each feature, and for each slice, intensity was plotted along the same line of pixels. A slope was found for every 5-pixel range along this line using a sliding window method,
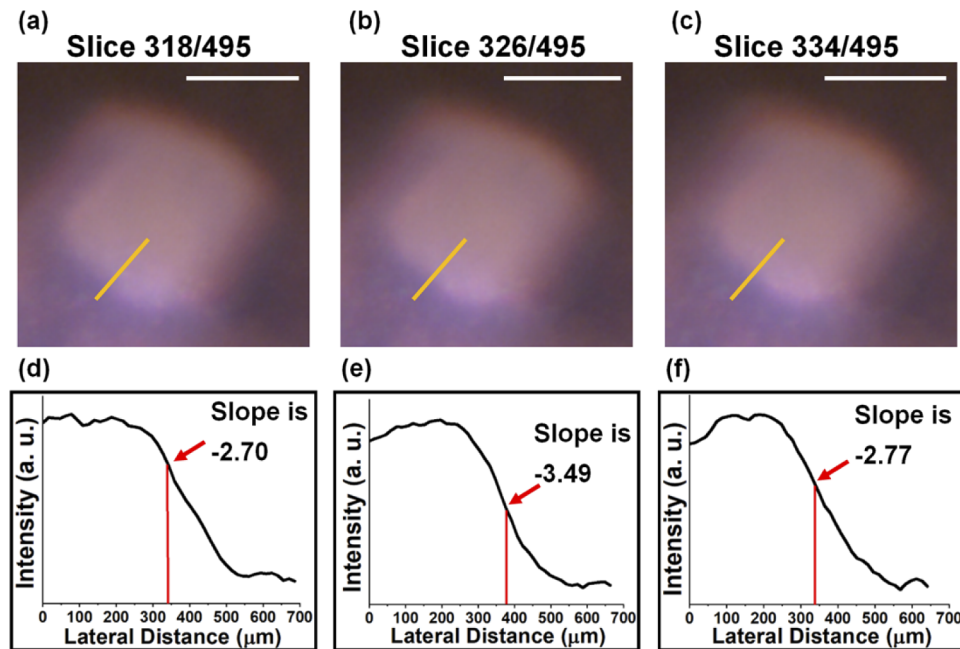

Fig. 15. Examples of determining focal plane for features of the heart model analyzed in Fig. 7. (a), (b), and (c) show images of the feature in (j) and (k) of Fig. 7 at different axial distances. (d), (e), and (f) show the corresponding intensity graphs along the yellow lines in the images and the maximum slope was found for each image. Scale bar: 1mm.

and the stack slice with the steepest slope was determined to be the slice where the edge of the feature was in focus. Using the data from the magnification program performed previously, the axial distance from the endoscope was found for the identified stack slice to determine the depth of the feature. This method was first applied to measure the axial distance to the flat face shown in Fig. 7(b) and Fig. 7(c) as the digital calipers used to provide a physical verification of the system's depth calculation were anchored on this face when measuring to the features shown in Fig. 7(e) and 7(f), Fig. 7(h) and 7(i), and Fig. 7(j) and 7(k). This process was performed to find the axial distance to all features shown in Fig. 7. The features in Fig. 7(e), Fig. 7(l), and Fig. 7(o) were also used to study the lateral resolution capabilities of the system. For these features, the steepest slope in the intensity plot was not only used to determine the axial distance to the feature, but the lateral location of the steepest slope in the intensity plot was used to determine the lateral location of the edge of a feature. Using this method, the lateral locations of two edges were found for the features marked in Fig. 7(f), Fig. 7(m), and Fig. 7(p), allowing for calculation of the width of these features.

Disclosures. The authors declare that there are no conflicts of interest to this article.

Data Availability. The code has been written in MATLAB (MathWorks). The latest version of the software is available in Ref. [31]. Other data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. R. Sinha, M. Sundaram, S. Raje, G. Rao, M. Sinha, and R. Sinha, "3D laparoscopy: technique and initial experience in 451 cases," Gynecol Surg **10**(2), 123–128 (2013).
2. A. Singh and R. Saraiya, "Three-dimensional endoscopy in sinus surgery," Curr Opin Otolaryngol Head Neck Surg **21**(1), 3–10 (2013).
3. D. Sahu, M. J. Mathew, and P. K. Reddy, "3D laparoscopy - help or hype; initial experience of a tertiary health centre," J. Clin. Diagn. Res. **8**(7), NC01-03 (2014).
4. G. Curro, G. La Malfa, A. Caizzone, V. Rampulla, and G. Navarra, "Three-dimensional (3D) versus two-dimensional (2D) laparoscopic bariatric surgery: a single-surgeon prospective randomized comparative study," OBES SURG **25**(11), 2120–2124 (2015).
5. M. J. Ali and M. N. Naik, "First intraoperative experience with three-dimensional (3D) high-definition (HD) nasal endoscopy for lacrimal surgeries," Eur Arch Otorhinolaryngol **274**(5), 2161–2164 (2017).
6. K. Vasudevan, H. Saad, and N. M. Oyesiku, "The role of three-dimensional endoscopy in pituitary adenoma surgery," Neurosurg Clin N Am **30**(4), 421–432 (2019).
7. K. Nomura, D. Kikuchi, M. Kaise, T. Iizuka, Y. Ochiai, Y. Suzuki, Y. Fukuma, M. Tanaka, Y. Okamoto, S. Yamashita, A. Matsui, T. Mitani, and S. Hoteya, "Comparison of 3D endoscopy and conventional 2D endoscopy in gastric endoscopic submucosal dissection: an ex vivo animal study," Surg Endosc **33**(12), 4164–4170 (2019).
8. J. D'Haens, K. Van Rompaey, T. Stadnik, P. Haentjens, K. Poppe, and B. Velkeniers, "Fully endoscopic transsphenoidal surgery for functioning pituitary adenomas: a retrospective comparison with traditional transsphenoidal microsurgery in the same institution," Surg Neurol **72**(4), 336–340 (2009).
9. B. Rotenberg, S. Tam, W. H. Ryu, and N. Duggal, "Microscopic versus endoscopic pituitary surgery: a systematic review," Laryngoscope **120**(7), 1292–1297 (2010).
10. T. Tokas, M. Avgeris, I. Leotsakos, U. Nagele, and A. S. Gozen, "Impact of three-dimensional vision in laparoscopic partial nephrectomy for renal tumors," Turk J. Urol. **47**(2), 144–150 (2021).
11. T. Kaltenbach, C. Leung, K. Wu, K. Yan, S. Friedland, and R. Soetikno, "Use of the colonoscope training model with the colonoscope 3D imaging probe improved trainee colonoscopy performance: a pilot study," Dig Dis Sci **56**(5), 1496–1502 (2011).
12. J. Spille, A. Wenners, U. von Hehn, N. Maass, U. Pecks, L. Mettler, and I. Alkatout, "2D versus 3D in laparoscopic surgery by beginners and experts: a randomized controlled trial on a pelvitrainer in objectively graded surgical steps," J Surg Educ **74**(5), 867–877 (2017).
13. R. Smith, K. Schwab, A. Day, T. Rockall, K. Ballard, M. Bailey, and I. Jourdan, "Effect of passive polarizing three-dimensional displays on surgical performance for experienced laparoscopic surgeons," Br J Surg **101**(11), 1453–1459 (2014).

14. C. Boulocher, E. Chereul, J. B. Langlois, M. Armenean, M. E. Duclos, E. Viguier, T. Roger, and E. Vignon, "Non-invasive in vivo quantification of the medial tibial cartilage thickness progression in an osteoarthritis rabbit model with quantitative 3D high resolution micro-MRI," Osteoarthritis Cartilage **15**(12), 1378–1387 (2007).

15. M. Kampschulte, C. R. Schneider, H. D. Litzlbauer, D. Tscholl, C. Schneider, C. Zeiner, G. A. Krombach, E. L. Ritman, R. M. Bohle, and A. C. Langheinrich, "Quantitative 3D micro-CT imaging of human lung tissue," Rofo **185**(09), 869–876 (2013).

16. L. Maier-Hein, P. Mountney, A. Bartoli, H. Elhawary, D. Elson, A. Groch, A. Kolb, M. Rodrigues, J. Sorger, S. Speidel, and D. Stoyanov, "Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery," Med Image Anal **17**(8), 974–996 (2013).

17. T. Nakagawa, T. Suzuki, Y. Hayashi, Y. Mizukusa, Y. Hatanaka, K. Ishida, T. Hara, H. Fujita, and T. Yamamoto, "Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair," J. Biomed. Opt. **13**(6), 064026 (2008).

18. J. J. Hyun, H. J. Chun, B. Keum, Y. S. Seo, Y. S. Kim, Y. T. Jeen, H. S. Lee, S. H. Um, C. D. Kim, H. S. Ryu, J. W. Lim, D. G. Woo, Y. J. Kim, and M. T. Lim, "Feasibility of obtaining quantitative 3-dimensional information using conventional endoscope: a pilot study," Clin Endosc **45**(3), 182–188 (2012).

19. R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," 12.

20. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light Field Microscopy," (ACM2006), pp. 924–934.

21. M. Levoy, Z. Zhang, and I. McDowall, "Recording and controlling the 4D light field in a microscope using microlens arrays," J. Microsc. **235**(2), 144–162 (2009).

22. A. Orth, M. Ploschner, E. R. Wilson, I. S. Maksymov, and B. C. Gibson, "Optical fiber bundles: ultra-slim light field imaging probes," Sci. Adv. **5**(4), eaav1555 (2019).

23. N. Bedard, T. Shope, A. Hoberman, M. A. Haralam, N. Shaikh, J. Kovačević, N. Balram, and I. Tošić, "Light field otoscope design for 3D in vivo imaging of the middle ear," Biomed. Opt. Express **8**(1), 260 (2017).

24. Y.-J. Wang, X. Shen, Y.-H. Lin, and B. Javidi, "Extended depth-of-field 3D endoscopy with synthetic aperture integral imaging using an electrically tunable focal-length liquid-crystal lens," Opt. Lett. **40**(15), 3564 (2015).

25. A. Hassanfiroozi, Y.-P. Huang, B. Javidi, and H.-P. D. Shieh, "Hexagonal liquid crystal lens array for 3D endoscopy," Opt. Express **23**(2), 971 (2015).

26. A. Hassanfiroozi, T.-H. Jen, Y.-P. Huang, and H.-P. D. Shieh, "Liquid crystal lens array for 3D endoscope application," in *SPIE Sensing Technology + Applications*, B. Javidi, J.-Y. Son, O. Matoba, M. Martínez-Corral, and A. Stern, eds. (2014), p. 91170E.

27. W.-T. Lin, C.-Y. Lin, V. R. Singh, and Y. Luo, "Speckle illumination holographic non-scanning fluorescence endoscopy," J. Biophotonics **11**(11), e201800010 (2018).

28. S. Zhu, P. Jin, R. Liang, and L. Gao, "Optical design and development of a snapshot light-field laryngoscope," Opt. Eng. **57**(2), 023110 (2018).

29. D. T. Moore, "Gradient-index optics: a review," Appl. Opt. **19**(7), 1035–1038 (1980).

30. C. Guo, T. Urner, and S. Jia, "3D light-field endoscopic imaging using a GRIN lens array," Appl. Phys. Lett. **116**(10), 101105 (2020).

31. T. Urner, A. Inman, B. Lapid, and S. Jia, "GLAM Software," https://github.com/ShuJiaLab/3D_endoscopy.

32. R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, P. Hanrahan, and D. Design, "Light Field Photography with a Hand-held Plenoptic Camera," Stanford Tech Report CTSR 2005-02, 11 (2005).

33. X. Sun, H. Ma, H. Ming, Z. Zheng, J. Yang, and J. Xie, "The measurement of refractive index profile and aberration of radial gradient index lens by using imaging method," Opt. Laser Technol. **36**(2), 163–166 (2004).

34. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," Nat. Methods **9**(7), 676–682 (2012).